ANSWERS AND EXPLANATIONS

Practice Examination 2

Section I: Multiple Choice Solutions

1. D

   The intercept is positive and approximately 9. The slope is negative, and the correlation coefficient is moderate, not strong.

2. C

   The given hypotheses were for a one-sided test. A one-sided hypothesis test with a significance level of 0.05 would correspond to a 90% confidence interval. Significant results would mean the sample results were significantly higher than the hypothesized value of 0.251. Therefore, 0.251 would fall below the lower limit of the interval.

3. E

   $\text{ME} = \frac{z^*\sigma}{\sqrt{n}}$. Solving for n yields $n = \left(\frac{z^*\sigma}{\text{ME}}\right)^2$.

   $0.5 = \frac{(1.96)(3)}{\sqrt{n}} \Rightarrow n = \left(\frac{(1.96)(3)}{0.5}\right)^2 \approx 138.2976$

   or 139 computer owners

4. D

   The purpose behind finding a confidence interval is to use a statistic to estimate a value for a parameter. If one is able to find the parameter, then there is no need to construct a confidence interval.

5. C

   Performing a two-proportion z-test with H_0: $p_M = p_F$ and H_a: $p_M \neq p_F$ yields a p-value of 0.324, insignificant results at any of the commonly accepted levels.

6. C

   Control, randomization, and replication are the key principles of experimental design. Randomization is used to reduce bias in results. Replication refers to the need to have an adequate sample size when conducting an experiment as well as the ability of another researcher to repeat your results. The placebo effect only affects humans. A placebo is not always necessary. You may be comparing two treatments. It is impossible to *eliminate all* lurking variables. (These words, "eliminate" and "all," should generally be avoided when describing experimental design.)

7. D

   A control group in this situation would add a baseline for comparison.

8. E

A χ^2 goodness-of-fit test should be performed. The conditions for conducting the test are met if we consider this shipment to be a random sample. Expected cell counts are each 100. The test statistic would be

$$\chi^2 = \frac{(89 - 100)^2}{100} + \frac{(95 - 100)^2}{100} + \frac{(106 - 100)^2}{100} + \frac{(110 - 100)^2}{100}$$

$= 2.82$. With three degrees of freedom, the p-value is 0.4202. Thus, the result is not significant at any of the commonly accepted significance levels.

9. A

The top 2% constitute the 98th percentile. The standard normal value for the 98th percentile is $z = 2.0537$. Using the formula for z-scores, the following equation can be set up:

$$2.0537 = \frac{41{,}293 - \mu}{7641}.$$

Solving for μ results in average concert attendance of 25,601.

10. B

These data are from a voluntary response sample, which is generally biased, and do not form a simple random sample of constituents.

11. E

Necessary conditions for a one-proportion z-interval are answer choices A, B, and C. All of these conditions are met.

A: We are told that the families were randomly chosen;

B: $50(10) = 500 < 1000$;

C: $50(0.60) = 30 > 10$ and $50(1 - 0.60) = 20 > 10$.

Answer choice D is *not* a necessary condition for proportions.

12. E

This is an example of a simple random sample. Each individual has the same probability of having his/her number selected, and every sample of size 5 has an equal probability of being selected.

13. B

We are told that the person was between the ages of 45 and 54. Of the 123 people in that category, 12 died as result of accidents. So, the probability is $12/123 \approx 0.0976$.

14. C

The data are paired by hotel. Since the rates are checked for the same hotels, we do not have two independent samples. Thus, a matched-pairs test is appropriate.

15. E

For a continuous probability density function, probability is associated with area. The sum of the areas exceeds 1. The area under the curve from $x = 0$ to $x = 8$ is $(0.1)(8) = 0.8$. The area under the curve from $x = 8$ to $x = 10$ is $(0.2)(2) = 0.4$. The sum of the two areas is $1.2 > 1$.

16. C

The range of average temperatures for July is 34.4 degrees, while the range of average temperatures for January is 66.2 degrees.

17. B

$$\frac{(\text{row total})(\text{column total})}{\text{table total}} = \frac{(71)(14)}{223} \approx 4.4574 \approx 4$$

18. D

This is the classic definition of double-blind. Neither the subjects nor the experimenters know who is in which group.

19. B

This is a geometric setting where $p = 0.450$ and $x = 3$. There are two failures ($q = 0.550$) before a success ($p = 0.450$).

20. B

The mean is not resistant to outliers; therefore, the mean will be affected by the outlying fruit with high IU content. On the other hand, the median is resistant to outliers and is therefore the better choice for a measure of center. The standard deviation, IQR, and range are measures of spread, not center.

21. A

$121/310 \approx 39\%$ is the proportion of sophomores who have a 3.0 to 4.0 GPA. The statement uses at least 3.0 and we must therefore count all who are 3.0 or more. This includes the cell counts 121 and 43 for a total of 164.

22. E

By definition, the coefficient of determination gives the percentage of the variability in the response variable explained by regression on the explanatory variable. Since we are only given a value for r^2, we cannot determine if the relationship between variables is positive or negative.

23. E

Once we remove the outlier for meat, it has the smallest range, IQR, and standard deviation.

24. B

Changing units will not affect the value of the correlation coefficient.

25. A

The distribution for grizzly bears is located to the right of that for black bears; therefore, grizzly bears tend to live longer. The distribution for grizzly bears is taller and thinner than that for black bears; therefore, there is less variability in the data for grizzly bears than black bears.

26. D

$$\text{ME} = z^*\sqrt{\frac{p(1-p)}{n}} = \frac{z^*\sqrt{p(1-p)}}{\sqrt{n}} \Rightarrow \sqrt{n} = \frac{z^*\sqrt{p(1-p)}}{\text{ME}}$$

$$n = \left(\frac{z^*\sqrt{p(1-p)}}{\text{ME}}\right)^2 = \left(\frac{1.645\sqrt{0.5(1-0.5)}}{0.03}\right)^2 \approx 751.674 \approx 752$$

27. A

By definition, two events are mutually exclusive if they cannot happen on the same outcome; therefore, the probability that both occur simultaneously is 0.

28. E

As the number of degrees of freedom *increases*, the *t*-distribution approaches the normal distribution.

29. A

All probabilities have values between 0 and 1, inclusive. The sum of the probabilities is 1.

30. E

Both the mean and the standard deviation will increase by 3.25%.

31. B

We wish to see if the average life expectancy in this village is more than 45 years of age.

32. E

A matched-pairs design will yield the best results for this experiment. Both socks will be exposed to the same usage and should give the most accurate results. Answer choice C was a completely randomized design. It is possible that one group would be much more active than the other, skewing the results. For answer choice D, it is possible that the two-week periods involved different activities, which could skew the results.

33. E

$$\begin{aligned} E(X) &= 2(0.3) + 4(0.2) + 6(0) + 8(0.4) + 10(0.1) \\ &= 0.6 + 0.8 + 0 + 3.2 + 1.0 \\ &= 5.6 \end{aligned}$$

$$\begin{aligned} \text{Var}(X) &= E(X-\mu)^2P(x) \\ &= (2-5.6)^2(0.3) + (4-5.6)^2(0.2) \\ &\quad + (6-5.6)^2(0) + (8-5.6)^2(0.4) + (10-5.6)^2(0.1) \\ &\approx 3.888 + 0.512 + 0 + 2.304 + 1.936 \\ &\approx 8.64 \end{aligned}$$

34. A

The residual with the least pattern was created with x and $\log y$ as the variables; therefore, the best fit is the equation $\log \hat{y} = 0.378x - 0.051$. When $x = 7$, $\log \hat{y} = 0.387(7) - 0.051 \approx 2.658$.

$$\hat{y} \approx 10^{2.658} \approx 454.988$$

35. B

The professors are stratified according to their status because professors of different status may have different feelings about tenure.

36. E

A z-score gives the number of standard deviations that a value falls below the mean for a negative score and above the mean for a positive score.

37. A

For the sampling distribution of sample proportions, $\mu_{\hat{p}} = p$ and $\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$. For this problem, $p = 0.135$ and $n = 75$. Standardize the endpoint values by subtracting this mean and dividing by this standard deviation.

38. E

The results can only be generalized for the population of individuals from which the sample was taken. In this case, that population is healthy teenagers.

39. A

This distribution is skewed right, so the mean will be larger than the median. The minimum is 0, and the maximum is 50. The lower quartile is between 0 and 5, and the upper quartile is 20. The median is between 5 and 10, closer to 10.

40. E

As the probability of a Type I error increases, power increases. As power increases, the probability of a Type II error decreases. Therefore, reducing power increases the probability of a Type II error, and reducing the probability of a Type I error increases the probability of a Type II error.